

ep



OIPE

## RAW SEQUENCE LISTING

DATE: 09/12/2003

PATENT APPLICATION: US/10/652,846

TIME: 12:39:30

Input Set : A:\D6020CIP4SEQ.txt

Output Set: N:\CRF4\09122003\J652846.raw

3 <110> APPLICANT: O'Brien, Timothy J.  
 4 Underwood, Lowell J.  
 6 <120> TITLE OF INVENTION: Extracellular Serine Protease  
 8 <130> FILE REFERENCE: D6020CIP3  
 W--> 10 <140> CURRENT APPLICATION NUMBER: US/10/652,846  
 C--> 10 <141> CURRENT FILING DATE: 2003-08-29  
 12 <150> PRIOR APPLICATION NUMBER: US 09/796,294  
 13 <151> PRIOR FILING DATE: 2001-02-28  
 15 <160> NUMBER OF SEQ ID NOS: 75  
 17 <210> SEQ ID NO: 1  
 18 <211> LENGTH: 144  
 19 <212> TYPE: PRT  
 20 <213> ORGANISM: unknown  
 22 <220> FEATURE:  
 23 <221> NAME/KEY: DOMAIN  
 24 <223> OTHER INFORMATION: Amino acid sequence of Protease m (Prom) catalytic  
 25 domain  
 27 <400> SEQUENCE: 1  
 28 Trp Val Leu Thr Ala Ala His Cys Lys Lys Pro Asn Leu Gln Val  
 29 5 10 15  
 30 Phe Leu Glu Lys His Asn Leu Arg Gln Arg Glu Ser Ser Gln Glu  
 31 20 25 30  
 32 Gln Ser Ser Val Val Arg Ala Val Ile His Pro Asp Tyr Asp Ala  
 33 35 40 45  
 34 Ala Ser His Asp Gln Asp Ile Met Leu Leu Arg Leu Ala Arg Pro  
 35 50 55 60  
 36 Ala Lys Leu Ser Glu Leu Ile Gln Pro Leu Pro Leu Glu Arg Asp  
 37 65 70 75  
 38 Cys Ser Ala Asn Thr Thr Ser Cys His Ile Leu Gly Trp Gly Lys  
 39 80 85 90  
 40 Thr Ala Asp Gly Asp Phe Pro Asp Thr Ile Gln Cys Ala Tyr Ile  
 41 95 100 105  
 42 His Leu Val Ser Arg Glu Glu Cys Glu His Ala Tyr Pro Gly Gln  
 43 110 115 120  
 44 Ile Thr Gln Asn Met Leu Cys Ala Gln Asp Glu Lys Tyr Gly Lys  
 45 125 130 135  
 46 Asp Ser Cys Gln Gly Asp Ser Gly Gly  
 47 140  
 49 <210> SEQ ID NO: 2  
 50 <211> LENGTH: 148  
 51 <212> TYPE: PRT  
 52 <213> ORGANISM: Homo sapiens  
 54 <220> FEATURE:

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/652,846

DATE: 09/12/2003

TIME: 12:39:30

Input Set : A:\D6020CIP4SEQ.txt

Output Set: N:\CRF4\09122003\J652846.raw

55 &lt;222&gt; LOCATION: DOMAIN

56 &lt;223&gt; OTHER INFORMATION: Amino acid sequence of Tadgl4 catalytic domain

58 &lt;400&gt; SEQUENCE: 2

```

59 Trp Val Val Thr Ala Ala His Cys Lys Lys Pro Lys Tyr Thr Val
60           5           10           15
61 Arg Leu Gly Asp His Ser Leu Gln Asn Lys Asp Gly Pro Glu Gln
62           20           25           30
63 Glu Ile Pro Val Val Gln Ser Ile Pro His Pro Cys Tyr Asn Ser
64           35           40           45
65 Ser Asp Val Glu Asp His Asn His Asp Leu Met Leu Leu Gln Leu
66           50           55           60
67 Arg Asp Gln Ala Ser Leu Gly Ser Lys Val Lys Pro Ile Ser Leu
68           65           70           75
69 Ala Asp His Cys Thr Gln Pro Gly Gln Asn Cys Thr Val Ser Gly
70           80           85           90
71 Trp Gly Thr Val Thr Ser Pro Arg Glu Asn Phe Pro Asp Thr Leu
72           95          100          105
73 Asn Cys Ala Glu Val Lys Ile Phe Pro Gln Lys Lys Cys Glu Asp
74          110          115          120
75 Ala Tyr Pro Gly Gln Ile Thr Asp Gly Met Val Cys Ala Gly Ser
76          125          130          135
77 Ser Lys Gly Ala Asp Thr Cys Gln Gly Asp Ser Gly Gly
78          140          145

```

80 &lt;210&gt; SEQ ID NO: 3

81 &lt;211&gt; LENGTH: 146

82 &lt;212&gt; TYPE: PRT

83 &lt;213&gt; ORGANISM: unknown

85 &lt;220&gt; FEATURE:

86 &lt;221&gt; NAME/KEY: DOMAIN

87 &lt;223&gt; OTHER INFORMATION: Amino acid sequence of trypsin like serine protease

88 (Try1) catalytic domain

90 &lt;400&gt; SEQUENCE: 3

```

91 Trp Val Val Ser Ala Gly His Cys Tyr Lys Ser Arg Ile Gln Val
92           5           10           15
93 Arg Leu Gly Glu His Asn Ile Glu Val Leu Glu Gly Asn Glu Gln
94           20           25           30
95 Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro Gln Tyr Asp Arg
96           35           40           45
97 Lys Thr Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Ser Ser Arg
98           50           55           60
99 Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser Leu Pro Thr Ala
100          65           70           75
101 Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly Asn
102          80           85           90
103 Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp Glu Leu Gln Cys Leu
104          95          100          105
105 Asp Ala Pro Val Leu Ser Gln Ala Lys Cys Glu Ala Ser Tyr Pro
106          110          115          120
107 Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly Phe Leu Glu Gly

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/652,846

DATE: 09/12/2003

TIME: 12:39:30

Input Set : A:\D6020CIP4SEQ.txt

Output Set: N:\CRF4\09122003\J652846.raw

```

108                               125                               130                               135
109 Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly
110                               140                               145
112 <210> SEQ ID NO: 4
113 <211> LENGTH: 144
114 <212> TYPE: PRT
115 <213> ORGANISM: unknown
117 <220> FEATURE:
118 <221> NAME/KEY: DOMAIN
119 <223> OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
120 enzyme (scce) catalytic domain
122 <400> SEQUENCE: 4
123 Trp Val Leu Thr Ala Ala His Cys Lys Met Asn Glu Tyr Thr Val
124                               5                               10                               15
125 His Leu Gly Ser Asp Thr Leu Gly Asp Arg Arg Ala Gln Arg Ile
126                               20                               25                               30
127 Lys Ala Ser Lys Ser Phe Arg His Pro Gly Tyr Ser Thr Gln Thr
128                               35                               40                               45
129 His Val Asn Asp Leu Met Leu Val Lys Leu Asn Ser Gln Ala Arg
130                               50                               55                               60
131 Leu Ser Ser Met Val Lys Lys Val Arg Leu Pro Ser Arg Cys Glu
132                               65                               70                               75
133 Pro Pro Gly Thr Thr Cys Thr Val Ser Gly Trp Gly Thr Thr Thr
134                               80                               85                               90
135 Ser Pro Asp Val Thr Phe Pro Ser Asp Leu Met Cys Val Asp Val
136                               95                               100                              105
137 Lys Leu Ile Ser Pro Gln Asp Cys Thr Lys Val Tyr Lys Asp Leu
138                               110                              115                              120
139 Leu Glu Asn Ser Met Leu Cys Ala Gly Ile Pro Asp Ser Lys Lys
140                               125                              130                              135
141 Asn Ala Cys Asn Gly Asp Ser Gly Gly
142                               140
144 <210> SEQ ID NO: 5
145 <211> LENGTH: 159
146 <212> TYPE: PRT
147 <213> ORGANISM: unknown
149 <220> FEATURE:
150 <221> NAME/KEY: DOMAIN
151 <223> OTHER INFORMATION: Amino acid sequence of hepsin (heps) catalytic domain
153 <400> SEQUENCE: 5
154 Trp Val Leu Thr Ala Ala His Cys Phe Pro Glu Arg Asn Arg Val
155                               5                               10                               15
156 Leu Ser Arg Trp Arg Val Phe Ala Gly Ala Val Ala Gln Ala Ser
157                               20                               25                               30
158 Pro His Gly Leu Gln Leu Gly Val Gln Ala Val Val Tyr His Gly
159                               35                               40                               45
160 Gly Tyr Leu Pro Phe Arg Asp Pro Asn Ser Glu Glu Asn Ser Asn
161                               50                               55                               60
163 Asp Ile Ala Leu Val His Leu Ser Ser Pro Leu Pro Leu Thr Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/652,846

DATE: 09/12/2003

TIME: 12:39:30

Input Set : A:\D6020CIP4SEQ.txt

Output Set: N:\CRF4\09122003\J652846.raw

```

164          65          70          75
165 Tyr Ile Gln Pro Val Cys Leu Pro Ala Ala Gly Gln Ala Leu Val
166          80          85          90
167 Asp Gly Lys Ile Cys Thr Val Thr Gly Trp Gly Asn Thr Gln Tyr
168          95          100          105
169 Tyr Gly Gln Gln Ala Gly Val Leu Gln Glu Ala Arg Val Pro Ile
170          110          115          120
171 Ile Ser Asn Asp Val Cys Asn Gly Ala Asp Phe Tyr Gly Asn Gln
172          125          130          135
173 Ile Lys Pro Lys Met Phe Cys Ala Gly Tyr Pro Glu Gly Gly Ile
174          140          145          150
175 Asp Ala Cys Gln Gly Asp Ser Gly Gly
176          155
178 <210> SEQ ID NO: 6
179 <211> LENGTH: 1360
180 <212> TYPE: DNA
181 <213> ORGANISM: Homo sapiens
183 <220> FEATURE:
184 <223> OTHER INFORMATION: Nucleotide sequence encoding Tumor Antigen Derived
185     Gene-14 (TADG-14) protein; nt 1344-1360 NCBI accession
186     #AA343629
188 <400> SEQUENCE: 6
189 ctgtagcagg cagagcttac caagtctctc cgaactcaaa tggagaagaat accttatgaa 60
190 tgtaagaatg taggggggtca tggcttgtaa ttacacaggt gtaaatgaaa ccatacctaga 120
191 ggattatgag gaatcctttc tatgtgattt tcaatcatag caagcaagaa aggctccagt 180
192 gtcaaggtag ttcagctctt acaggatata aaacagtgca tacttgagag aaaaaactta 240
193 gatctgagtg atggaatgtg aagcaaattc ttcaaaatca gtagacattt cttggacata 300
194 aaacacagat gaggaaaggg cttcaaatta gaagttacgt aatcaccatc agaaagttca 360
195 tgttttgtaa attctgttac tagaaatgta ggaaattcag gtatagcttt gaatcccaat 420
196 tacacattgg tcagtgggaa aactaagggc ctccaacagg caaattcagg gaggataggt 480
197 ttcagggaat gccctggatt ctggaagacc tcacatggg acgccccga cctcgtgcgg 540
198 ccaagacgtg gatgttcctg ctcttgctgg ggggagcctg ggcaggacac tccagggcac 600
199 aggaggacaa ggtgctgggg ggtcatgagt gccaaccca ttcgcagcct tggcaggcgg 660
200 ccttgttcca gggccagcaa ctactctgtg gcggtgtcct tgtaggtggc aactgggtcc 720
201 ttacagctgc cactgtaaa aaaccgaaat acacagtagc cctgggagac cacagcctac 780
202 agaataaaga tggcccagag caagaaatac ctgtggttca gtccatcca caccctgct 840
203 acaacagcag cgatgtggag gaccacaacc atgatctgat gcttcttcaa ctgcgtgacc 900
204 aggcatccct ggggtccaaa gtgaagccca tcagcctggc agatcattgc accagcctg 960
205 gccagaagtg caccgtctca ggctggggca ctgtcaccag tccccgagag aattttcctg 1020
206 acactctcaa ctgtgcagaa gtaaaaatct ttcccagaa gaagtgtgag gatgcttacc 1080
207 cggggcagat cacagatggc atggtctgtg caggcagcag caaaggggct gacacgtgcc 1140
208 agggcgattc tggaggcccc ctggtgtgtg atggtgcact ccagggcatc acatcctggg 1200
209 gctcagaccc ctgtgggagg tccgacaaac ctggcgctta taccaacatc tgccgctacc 1260
210 tggactggat caagaagatc ataggcagca agggctgatt ctaggataag cactagatct 1320
211 cccttaataa actcacaact ctctgaaaaa aaaaaaaaaa 1360
213 <210> SEQ ID NO: 7
214 <211> LENGTH: 260
215 <212> TYPE: PRT
216 <213> ORGANISM: Homo sapiens

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/652,846

DATE: 09/12/2003

TIME: 12:39:30

Input Set : A:\D6020CIP4SEQ.txt

Output Set: N:\CRF4\09122003\J652846.raw

218 &lt;220&gt; FEATURE:

219 &lt;223&gt; OTHER INFORMATION: Amino acid sequence of TADG-14 protein

221 &lt;400&gt; SEQUENCE: 7

```

222 Met Gly Arg Pro Arg Pro Arg Ala Ala Lys Thr Trp Met Phe Leu
223           5           10           15
224 Leu Leu Leu Gly Gly Ala Trp Ala Gly His Ser Arg Ala Gln Glu
225           20           25           30
226 Asp Lys Val Leu Gly Gly His Glu Cys Gln Pro His Ser Gln Pro
227           35           40           45
228 Trp Gln Ala Ala Leu Phe Gln Gly Gln Gln Leu Leu Cys Gly Gly
229           50           55           60
230 Val Leu Val Gly Gly Asn Trp Val Leu Thr Ala Ala His Cys Lys
231           65           70           75
232 Lys Pro Lys Tyr Thr Val Arg Leu Gly Asp His Ser Leu Gln Asn
233           80           85           90
234 Lys Asp Gly Pro Glu Gln Glu Ile Pro Val Val Gln Ser Ile Pro
235           95          100          105
236 His Pro Cys Tyr Asn Ser Ser Asp Val Glu Asp His Asn His Asp
237          110          115          120
238 Leu Met Leu Leu Gln Leu Arg Asp Gln Ala Ser Leu Gly Ser Lys
239          125          130          135
240 Val Lys Pro Ile Ser Leu Ala Asp His Cys Thr Gln Pro Gly Gln
241          140          145          150
242 Lys Cys Thr Val Ser Gly Trp Gly Thr Val Thr Ser Pro Arg Glu
243          155          160          165
244 Asn Phe Pro Asp Thr Leu Asn Cys Ala Glu Val Lys Ile Phe Pro
245          170          175          180
246 Gln Lys Lys Cys Glu Asp Ala Tyr Pro Gly Gln Ile Thr Asp Gly
247          185          190          195
248 Met Val Cys Ala Gly Ser Ser Lys Gly Ala Asp Thr Cys Gln Gly
249          200          205          210
250 Asp Ser Gly Gly Pro Leu Val Cys Asp Gly Ala Leu Gln Gly Ile
251          215          220          225
252 Thr Ser Trp Gly Ser Asp Pro Cys Gly Arg Ser Asp Lys Pro Gly
253          230          235          240
254 Val Tyr Thr Asn Ile Cys Arg Tyr Leu Asp Trp Ile Lys Lys Ile
255          245          250          255
256 Ile Gly Ser Lys Gly
257          260

```

259 &lt;210&gt; SEQ ID NO: 8

260 &lt;211&gt; LENGTH: 260

261 &lt;212&gt; TYPE: PRT

262 &lt;213&gt; ORGANISM: Mus sp.

264 &lt;220&gt; FEATURE:

265 <223> OTHER INFORMATION: Amino acid sequence of mouse neuropsin homologous  
 266 to TADG-14; accession no. D30785

268 &lt;400&gt; SEQUENCE: 8

```

269 Met Gly Arg Pro Pro Pro Cys Ala Ile Gln Pro Trp Ile Leu Leu
270           5           10           15

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/652,846

DATE: 09/12/2003

TIME: 12:39:31

Input Set : A:\D6020CIP4SEQ.txt

Output Set: N:\CRF4\09122003\J652846.raw

L:10 M:282 W: Numeric Field Identifier Missing, <140> CURRENT APPLICATION NUMBER: is Added.  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:453 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:12  
L:465 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:13